

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-69537

(43) 公開日 平成10年(1998) 3月10日

(51) Int.Cl. <sup>8</sup>	識別記号	序内整理番号	F I	技術表示箇所
G 0 6 T	1/00		G 0 6 F 15/66	4 7 0 J
	7/00		15/70	3 3 0 P
				4 6 0 A

審査請求 有 請求項の数10 O L (全 14 頁)

(21) 出願番号 特願平8-226879

(22) 出願日 平成8年(1996) 8月28日

(71) 出願人 000004237

日本電気株式会社  
東京都港区芝五丁目7番1号

(72) 発明者 中島 昇

東京都港区芝五丁目7番1号 日本電気株  
式会社内

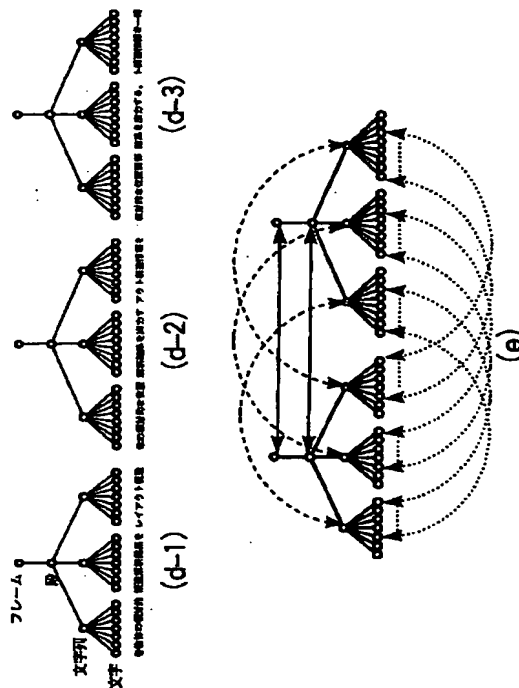
(74) 代理人 弁理士 後藤 洋介 (外2名)

(54) 【発明の名称】 画像合成方法及び画像合成装置

(57) 【要約】

【課題】 低解像度の画像入力デバイスを用いて、特別な機械的走査機構を用いずに広範囲の高解像度画像を取得する。

【解決手段】 低解像度の画像入力デバイスを手動で走査して、画像入力部10から複数枚のフレーム画像を取得し、入力された複数枚の前記フレーム画像のそれぞれに対して、レイアウト構造解析部31においてレイアウト解析処理を行う。レイアウト構造対応づけ部41では、前記レイアウト構造をフレーム画像間で対応づけ、これに基づきフレーム画像間の相対的なずれを位置ずれ量として検出する。各物体ごとに画素の大きさ未満の位置ずれ量を検出し、超解像処理部80において超解像処理により高解像度の超解像画像を取得する。画像合成部60において、レイアウト構造に基づき、超解像画像をつなぎあわせることで広範囲にわたる高解像度画像を生成する。



## 【特許請求の範囲】

【請求項1】 複数の画像からひとつの画像を合成する画像合成方法において、  
前記複数枚のフレーム画像のそれぞれに対して、フレーム画像内の各物体の検出及び検出された各物体について物体の特徴と物体相互の位置関係からなる構造情報の抽出を行う構造解析段階と、  
前記構造情報を前記複数枚のフレーム画像の間で対応づけ、これに基づきフレーム画像間の相対的な位置ずれ量を検出する構造情報対応づけ段階と、  
前記複数枚のフレーム画像を前記位置ずれ量に基づいて繋ぎ合わせる画像合成段階とを含むことを特徴とする画像合成方法。

【請求項2】 請求項1記載の画像合成方法において、前記フレーム画像は文書の画像であり、前記構造情報として段組み、文字列及び文字からなる文書のレイアウト構造を抽出することを特徴とする画像合成方法。

【請求項3】 請求項1記載の画像合成方法において、前記複数枚のフレーム画像の中の同一物体に対して、前記構造情報を基に一画素の大きさ以下の精度の位置ずれ量であるサブピクセル位置ずれ量を検出する位置ずれ残差検出段階と、  
前記サブピクセル位置ずれ量に基づき、画素間を内挿するように画像をずらして重ねあわせた超解像画像を生成する超解像処理段階と、  
前記構造情報に基づき前記超解像画像を画像上に配置して画像を生成する超解像画像合成段階とを含むことを特徴とする画像合成方法。

【請求項4】 請求項2記載の画像合成方法において、前記複数枚のフレーム画像の中の同一物体について、前記構造情報を基に前記サブピクセル位置ずれ量を検出する位置ずれ残差検出段階と、  
前記サブピクセル位置ずれ量に基づき、画素間を内挿するように画像をずらして重ねあわせた超解像画像を各物体について生成する超解像処理段階と、  
前記構造情報に基づき前記超解像画像を画像上に配置して画像を生成する超解像画像合成段階とを含むことを特徴とする画像合成方法。

【請求項5】 請求項4記載の画像合成方法において、前記超解像画像を文字単位に生成した超解像文字画像を出力する超解像文字画像出力段階と、  
前記超解像文字画像を文字コードに変換する個別文字認識段階と、  
前記構造情報と前記位置ずれ量に基づき、前記文字コードを文字の記載順にしたがって配置し、文書の認識結果を再構成する文字認識結果結合段階とを含むことを特徴とする画像合成方法。

【請求項6】 複数の画像からひとつの画像を合成する画像合成装置において、  
入力対象に対して画像入力デバイスを走査して複数枚の

フレーム画像を取得する画像入力手段と、  
入力された前記複数枚のフレーム画像を記憶する画像記憶手段と、  
前記複数枚のフレーム画像のそれぞれに対して、フレーム画像内の各物体の検出及び検出された各物体について前記構造情報の抽出を行う構造解析手段と、  
前記構造解析手段により前記複数枚のフレーム画像から抽出された前記構造情報を記憶する構造情報記憶手段と、

10 前記構造情報記憶手段に記憶された前記構造情報を前記複数枚のフレーム画像の間で対応づけ、これに基づき各フレーム画像間の相対的な位置ずれ量を検出し、前記構造情報記憶手段に出力する構造情報対応づけ手段と、  
前記画像記憶手段に記憶されている前記複数枚のフレーム画像を、前記構造情報記憶手段に記憶されたフレーム画像間の位置ずれ量に基づき、フレーム画像を繋ぎ合わせる画像合成手段とを含むことを特徴とする画像合成装置。

【請求項7】 請求項6記載の画像合成装置において、  
20 入力対象は文書であり、前記構造情報として段組み、文字列及び文字からなる文書のレイアウト構造を抽出することを特徴とする画像合成装置。

【請求項8】 請求項6記載の画像合成装置において、前記複数枚のフレーム画像の中の同一物体に対して、前記構造情報を基に前記サブピクセル位置ずれ量を検出する位置ずれ残差検出手段と、前記サブピクセル位置ずれ量に基づき、画素間を内挿するように画像をずらして重ねあわせた超解像画像を生成する超解像処理手段と、  
30 前記構造情報に基づき前記超解像画像を画像上に配置して画像を生成する超解像画像合成手段とを含むことを特徴とする画像合成装置。

【請求項9】 請求項7記載の画像合成装置において、前記複数枚のフレーム画像の中の同一物体について、前記構造情報を基に前記サブピクセル位置ずれ量を検出する位置ずれ残差検出手段と、  
前記サブピクセル位置ずれ量に基づき、画素間を内挿するように画像をずらして重ねあわせた超解像画像を各物体について生成する超解像処理手段と、  
40 前記構造情報に基づき前記超解像画像を画像上に配置して画像を生成する超解像画像合成手段とを含むことを特徴とする画像合成装置。

【請求項10】 請求項9記載の画像合成装置において、  
前記超解像文字画像を出力する超解像文字画像出力手段と、  
前記超解像文字画像を文字コードに変換する個別文字認識手段と、  
前記構造情報と前記位置ずれ量に基づき、前記文字コードを文字の記載順にしたがって配置し、文書の認識結果

を再構成する文字認識結果結合手段とを含むことを特徴とする画像合成装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は画素数の少ない、例えばCCDカメラのような画像入力装置を利用して高解像度の画像を取得する画像取得方法に関し、特に文書認識システムにおける文字画像の高解像度化を行う方法に関する。

【0002】

【従来の技術】高解像度で広範囲にわたる画像を生成するために、複数の画像からひとつの画像を合成する画像合成方法が用いられている。従来の高解像度画像取得方法としては、(1)高解像度の部分画像を撮像範囲をオーバーラップさせて複数枚入力し、重複部分から画像間の位置ずれを検出し、画像を接続する方法、(2)低解像度の全体画像を複数枚入力し、各画像間の位置ずれを検出することで、画素を内挿する方法が提案されている。

【0003】前記(1)の一例が、信学技報IE81-17に記載されている。この論文では、CCDカメラを手動で移動させながら、入力対象を部分ごとに分けて接写することで高解像度の部分画像を複数枚入力し、これらを接続して広範囲にわたる高解像度画像を取得する方法を提案している。

【0004】具体的には、重複部分を持たせて撮影した部分画像間の位置のずれを検出して、部分画像を接続する。ここで、部分画像間の位置ずれ量は、画像を平行移動しながら画像を重ねあわせて残差の計算を行い、残差が最小になったときの移動量として検出する。

【0005】前記(2)の一例が、1995年アイ・シー・アイ・エー・ピー予稿、387ページから392ページ(Proceedings of 8th ICIAP, pp. 387-392)に記載されている。この論文では、入力対象全体を低解像度の画像として複数枚入力し、複数の画像を用いて画素間の輝度値を内挿することで、広範囲にわたる高解像度画像を再構成する方法が提案されている。

【0006】複数枚の画像を用いて画素間の輝度値を内挿するためには、各低解像度画像どうしの位置ずれを画素単位未満(以降で、サブピクセルと呼ぶ)の精度で検出する必要がある。この論文の方法では位置ずれを検出する方法として、画像の部分領域どうしの相関値が最大になるように部分領域の画像を移動、変形して局所的な対応づけを行う。このとき、局所領域の変形は、最急降下法を用いて逐次的に最適な変形を探索する。

【0007】

【発明が解決しようとする課題】前記の従来の方法

(1)には下記の2つの問題がある。

【0008】第1に、カメラの移動方向が未知であるため、残差値が最小になる移動量を全ての可能な重ねあわ

せ方について全探索しなければならない。このためには膨大な計算量が必要となる。

【0009】第2に、手動走査により得られた部分画像を接続して広範囲にわたる画像を取得するが、全ての部分画像が所望の解像度を保って高解像度に撮影されている必要がある。このためには、CCDカメラと対象間の距離を一定距離以内に保った手動走査を行わなくてはならない。これにより、手軽で自由度が大きいという手動走査の利点が損なわれる。

10 【0010】前記の従来の方法(2)は下記の3つの問題がある。

【0011】第1に、移動量の全探索を避けるために、相関値が大きくなるように逐次局所画像を変形して画像間の位置ずれを検出するが、対象の画像が空間的な周期を持つ場合、正しい位置ずれ量が算出される保証がない。たとえば、文書画像を対象とする場合、文字列、文字、文字のストロークの周期性により、位置ずれ量が誤って検出される可能性がある。

20 【0012】第2に、画像全面にテクスチャの存在しないような画像にこの手法を適用する場合に検出される位置ずれ量の精度が劣化しやすい。例えば、文書画像のように、行間、段間等画像領域のほとんどを背景部の白画素が占めている場合には、残差、相関の値に極値が現れず位置ずれの検出ができない。

【0013】第3に、複数の画像を用いて輝度値を内挿することで解像度の改善を行うが、2枚の画像の移動量が画素間隔の $1/2$ の場合には解像度を2倍に上げられる可能性がある。しかし、2枚の画像の位置ずれ量が0の場合には解像度は全く改善されない。CCDカメラを手動により移動した場合には、画素未満の位置ずれ量を随意に制御することが不可能であり、解像度の改善の度合いは不安定となる。

【0014】本発明の目的は、手動によりラフなCCDカメラ走査を行った場合でも広範囲にわたる高解像度画像を手軽に取得できる高解像度画像取得方式を提供することにある。

【0015】また、本発明の他の目的は、例えば文書画像のような高周波のテクスチャが局在するような画像においても、高速かつ安定に画像の対応づけが行え、広範囲かつ高解像度な画像を取得可能な高解像度画像取得方式を提供することにある。

【0016】さらに、本発明によれば、内挿による解像度の改善処理の効果の不安定さを補い、安定に解像度の向上を実現できる。

【0017】

【課題を解決するための手段】第1に、本発明の高解像度画像取得方式は、まず画像の構造解析を行い、構造情報の対応づけを行うことで、フレーム画像間の相対的な位置ずれを検出し、画像の接続を行う。

【0018】より具体的には、まず構造解析処理を行い

画像内の物体の特徴と物体間の位置関係を抽出する構造解析部（図 1 の 30）と、各フレーム画像から抽出された構造情報に対応づけ、フレーム画像間の相対的な位置関係を算出する構造情報対応づけ部（図 1 の 40）とを有する。

【0019】第 2 に、本発明の高解像度画像取得方式は、文書画像に適用する場合には、構造情報としてレイアウト構造を用いる。

【0020】より具体的には、文書画像内の段、文字列、文字といった物体とその配置の情報からなるレイアウト構造を抽出するレイアウト解析部（図 3 の 31）と、各フレーム画像から抽出されたレイアウト構造に対応づけ、フレーム画像間の相対的な位置関係を算出するレイアウト構造対応づけ部（図 3 の 41）とを有する。

【0021】第 3 に、本発明の高解像度画像取得方式は、対応づけられた構造情報に基づき、画像の接続を行うとともに、画像を構成する物体ごとに画像の解像度を改善する。

【0022】より具体的には、構造情報に基づき各フレーム画像から抽出した物体間の画素の大きさ未満の位置ずれ量であるサブピクセル位置ずれ量を算出する位置ずれ残差検出部（図 5 の 60）と、サブピクセル位置ずれ量を考慮して画素間を内挿することで画像の解像度を向上させる超解像処理部（図 5 の 70）とを有する。

【0023】第 4 に、本発明の高解像度画像取得方式は、対応づけられたレイアウト構造情報に基づき、文書画像の接続を行うとともに、文書画像を構成する物体ごとに画像の解像度を改善する。

【0024】より具体的には、レイアウト構造に基づき各フレーム画像から抽出した物体間の画素の大きさ未満の位置ずれ量であるサブピクセル位置ずれ量を算出する位置ずれ残差検出部（図 7 の 60）と、サブピクセル位置ずれ量を考慮して画素間を内挿することで画像の解像度を向上させる超解像処理部（図 5 の 70）とを有する。

【0025】第 5 に、本発明の高解像度画像取得方式は、解像度の改善された物体の画像ごとに文字を認識し、接続されたレイアウト構造に基づき、文字認識結果を並び替えることで文書全体の認識を行う。

【0026】より具体的には、解像度の改善された文字画像を入力し文字コードに変換する個別文字認識部（図 8 の 90）と、統合されたレイアウト構造に基づき、文字コードを文書での記載の順序に並び替える文字認識結果結合部（図 8 の 81）とを有する。

【0027】構造情報対応づけ部は、構造解析部において抽出された各フレーム画像ごとの構造情報どうしの対応関係を抽出し、対応関係をもとにフレーム画像どうしの位置ずれ量を算出する。

【0028】レイアウト構造対応づけ部はレイアウト解析部において抽出された各フレーム画像ごとのレイアウト

構造どうしの対応関係を抽出し、対応関係をもとにフレーム画像どうしの位置ずれ量を算出する。

【0029】超解像処理部は対応づけられた構造情報に基づき、画像の接続を行うと同時に、画像を構成する物体ごとに画像の解像度を改善する。

【0030】個別文字認識部では、解像度の改善された文字画像が入力され、対応する文字コードが出力される。

【0031】文字認識結果結合部はレイアウト構造対応づけ部において対応づけられたレイアウト構造に基づき、文字コードを文書の読み取り順に並び替えることで、文書全体にわたる認識結果を得る。

【0032】

【発明の実施の形態】本発明の第 1 の実施の形態について説明する。

【0033】図 1 は、本発明の第 1 の実施の形態の構成を示すブロック図である。本発明の第 1 の実施の形態は、画像入力デバイスを手動により走査することで得られる画像信号をデジタル化して複数枚の多値のフレーム画像として計算機に入力する画像入力部 10、計算機に入力されたフレーム画像を記憶する画像記憶部 20、画像に含まれる物体の構造情報を解析する構造解析部 30、各フレーム画像から抽出された構造情報どうしの対応づけを行う構造情報対応づけ部 40、構造解析の結果およびフレーム間で構造情報を構造情報記憶部 50 及び対応づけられた構造情報をもとに複数のフレーム画像を合成する画像合成部 60 とからなる。

【0034】図 1 及び図 2 を参照して、本発明の第 1 の実施の形態の動作について説明する。

【0035】図 2（a）のような画像入力対象物に対して、例えば同図に示した点線のように画像入力デバイスを移動して、画像入力部 10 より分割して入力する。

【0036】入力された画像はそれぞれ、図 2（b-1）（b-2）のようになり、画像記憶部 20 に記録される。

【0037】同時に、入力画像は構造解析部 30 に入力される。構造解析部 30 は、画像の構造情報を解析する。例えば、画像を 2 値化して、黒画素連結成分を抽出し物体を分離して抽出する。

【0038】抽出された物体の個々の特徴および物体間の位置関係を構造情報として、構造情報記憶部 50 に出し記憶する。

【0039】構造情報対応づけ部 40 では構造情報記憶部 50 に記憶された各フレーム画像の構造情報を入力し、構造情報を用いて、各フレーム内の対応する物体どうしを対応づける。

【0040】対応づけられた 2 枚のフレーム画像の構造情報は、対応関係から統合され、統合された構造情報は構造情報記憶部 50 に出力される。

【0041】画像合成部 60 は、統合された構造情報を

もとにフレーム画像を画像記憶部20から取り出して接合する。

【0042】次に、図2を用いて、撮影された物体の2枚の画像が本発明の第1の実施の形態によって合成される過程を説明する。

【0043】図2を参照すると、図2(a)のような物体が同図(b-1)(b-2)のように分割して撮影され、画像入力部10より入力される。

【0044】構造解析部30では、同図(c-1)(c-2)に示したように構造情報が抽出される。構造情報として、例えば、個々の物体の形状情報、2つの物体の相互の位置関係を用いる。

【0045】図2の例では、個々の物体の形状情報として、「○」、「△」、「□」の各カテゴリを用いる。形状情報の抽出は、画像上に出現すると考えられる物体のテンプレート「○」、「△」、「□」等を用意しておき、各テンプレートを回転、拡大縮小しながら画像とマッチングし、最もよくマッチするテンプレートのカテゴリを物体の形状とする。

【0046】形状情報は、2つの画像に現れる同一物体を同一であると識別できればよく、必ずしも物体とテンプレートが完全に一致する必要はないので、この例のように「○」、「△」、「□」といった代表的な形状のテンプレートがあれば良い。例えば、一方の画像で歪んだ五角形がテンプレート「△」を回転、拡大縮小したものと最も良くマッチしたとしても、もう一方の画像の同じ五角形がテンプレート「△」を同様に回転、拡大縮小したものと認識出来ればよく、五角形のテンプレートを用意する必要はない。

【0047】また、2つの物体間の相互の位置関係としては、両物体を結ぶ線分の長さ・線分の向き、両物体のモーメントを用いる。

【0048】また、物体の形状を検出する際に適用した回転量を回転角として各物体ごとに記憶する。物体内部に含まれる画素の数を面積として記憶する。各物体の重心点をもとめ、フレーム画像内に含まれる全ての物体のうちの2物体の全ての組み合わせについて重心点間を結ぶベクトルを記憶する。最後に、各フレーム画像の原点を基準とした各物体の重心点の座標をベクトルとして記憶する。

【0049】以上の各形状情報、物体の位置関係、回転角、面積、重心点間のベクトル、重心点のベクトルからなる情報を構造情報とし、構造情報記憶部50に出力する。図2を参照すると、物体は1フレーム目で左から「△」101、「□」102、「○」103の3個、2フレーム目では「□」104、「○」105、「□」106の3個がそれぞれ検出され、各構造情報は同図(c-1)(c-2)および図3のようになる。

【0050】次に、構造情報対応づけ部40において、一方のフレーム画像上に存在する物体のなかから2つの

物体の組を選び、他方の画像上に2つの物体と同じ形状の物体が含まれる場合には、両フレーム画像から抽出した物体の組の相対位置の絶対値の差を計算する。

【0051】フレーム画像内に含まれる全ての2物体の組について、同様の処理を行い、相対位置の絶対値の差が最小および2番目に最小となる物体の組をそれぞれ検索する。各組に関して、フレーム画像の間の相対的な平行移動量、回転量を計算する。

【0052】図2(c-1)(c-2)を参照すると、(c-1)のフレーム画像における物体102の形状および面積は、(c-2)のフレーム画像の物体104と一致する。また、(c-1)のフレーム画像における物体103の形状および面積は、(c-2)のフレーム画像の物体105と一致する。

【0053】さらに、(c-1)のフレーム画像における物体102と物体103を結ぶベクトル $v_{102}$ の長さは、(c-2)のフレーム画像の物体104と物体103を結ぶベクトル $v_{102}$ の長さとも一致する。

【0054】また、前記の対応関係に位置・回転に関して矛盾がないため、物体102と物体104、および物体103と物体105が一致すると判定する。フレーム画像間の位置ずれ量 $s$ は $(x_{104} - x_{102})$ 、または $(x_{105} - x_{103})$ となり、この値は構造情報記憶部50に出力される。量子化誤差の影響も考慮して、例えば両者の平均をとってもいい。この位置ずれ量に基づき、(c-2)の2枚目のフレーム画像を(c-1)の1枚目のフレーム画像の座標系に変換して、構造情報記憶部50に出力する。

【0055】本実施例において各物体の特徴量は図4のように変換される。ここで、一致する物体は1つの物体として記憶される。

【0056】次に、画像合成部60は2枚のフレーム画像および位置ずれ量 $s$ を入力し、位置ずれ量を考慮して2枚のフレーム画像を接続する。この結果、図2(d)の接続画像が得られる。

【0057】続いて、3枚目以降のフレーム画像が入力される場合は、3枚目に関して構造解析部30において、前記と同様な構造解析処理が施され、構造解析結果が構造情報記憶部50に出力される。構造情報対応づけ部40ではこの構造情報と、1枚目と2枚目を統合して得られた構造情報の間で対応づけを行い、画像合成部60でのフレーム画像の接続処理を行う。4枚目以降も同様に逐次接続処理を行っていく。

【0058】以上のように、本発明の第1の実施の形態によれば、入力された各フレーム画像の構造を解析し構造情報どうしの対応づけを行うため、フレーム画像の位置ずれを検出するための相関計算の回数を削減でき、位置ずれ量の検出精度の向上も実現できる。

【0059】本発明の第2の実施の形態について図面を参照して詳細に説明する。

【0060】第2の実施の形態では、入力する対象を文書に限定する。図5を参照すると、本発明の第2の実施の形態の構成は、第1の実施の形態における構造解析部30はレイアウト構造解析部31に、第1の実施の形態の構造情報対応づけ部40はレイアウト構造対応づけ部41に、第1の実施の形態における構造情報記憶部50はレイアウト構造記憶部51に置換される点で異なる。

【0061】レイアウト構造解析部31は、分割して入力される文書のフレーム画像に、段・文字列・文字等がどの位置にどのような方向で記述されているかを解析し、例えば、フレーム—段—文字列—文字というような階層構造で表現されるレイアウト構造を生成し、これをレイアウト構造記憶部51に出力する。

【0062】レイアウト構造対応づけ部はレイアウト構造記憶部51に記憶されている複数のフレーム画像から生成されたレイアウト構造内の一致する物体どうしを対応づける。ここで、物体とはレイアウト構造の各階層に相当するフレーム、段、文字列、文字である。

【0063】各フレーム画像で生成されたレイアウト構造を1つのレイアウト構造に統合し、構造情報記憶部50に出力する。

【0064】また、フレーム画像間の位置ずれ量をレイアウト構造の対応関係から算出し、構造情報記憶部50に出力する。

【0065】レイアウト構造記憶部51はレイアウト構造解析部31の出力するレイアウト構造情報を一時記憶し、レイアウト構造対応づけ部に出力する。

【0066】また、レイアウト構造対応づけ部41において複数のフレーム画像から生成されたレイアウト構造を統合した結果のレイアウト構造を入力し記憶する。

【0067】同時に、レイアウト構造対応づけ部41の出力するフレーム画像間の位置ずれ情報を入力し記憶する。

【0068】フレーム画像間の位置ずれ情報は適宜画像合成部60に出力される。

【0069】本発明の第2の実施の形態の動作を図5を参照して詳細に説明する。

【0070】レイアウト構造解析部31は、画像入力部10において入力された文書を部分的に撮影して得られたフレーム画像に対してレイアウト解析処理を施し、レイアウト構造情報を得る。レイアウト解析処理の方法としては、例えば、電子情報通信学会論文誌、V o l . J 7 4 - D - I I , N o . 4 , 4 9 1 ページから4 9 9 ページに記載されている方法を用いることができる。

【0071】解析処理により抽出されたレイアウト構造情報は、レイアウト構造記憶部51に出力され記憶される。

【0072】レイアウト構造対応づけ部41はレイアウト構造記憶部51に記憶された複数のフレーム画像から生成されたレイアウト構造情報を入力し、レイアウト構

造の対応づけを行う。

【0073】対応づけられたレイアウト構造は統合され、レイアウト構造記憶部51に出力される。このとき、レイアウト構造の対応関係から、フレーム画像間の位置ずれ量を算出し、レイアウト構造記憶部51に出力する。

【0074】画像合成部60は画像記憶部20に記憶された複数枚のフレーム画像と、レイアウト構造記憶部51に記憶された位置ずれ量を入力し、フレーム画像を重ね合わせて一枚の画像に合成する。

【0075】次に、図6を用いて、3枚の文書の画像が本発明の第2の実施の形態によって合成される過程を説明する。

【0076】図6(a)のような文書を同図矢印で示したように撮像デバイスを移動させながら、(b-1)～(b-3)のように分割して撮影された部分的なフレーム画像が画像入力部10より入力される。

【0077】レイアウト構造解析部31では、入力されたフレーム画像に対してレイアウト解析処理が施され、同図(c-1)～(c-3)に示したように領域分割され、各分割領域の属性、各分割領域間の関係が抽出される。具体的には、1枚のフレーム画像が複数の「段」に、1つの「段」が複数の「文字列」に、1つの「文字列」が複数の「文字」に分割され、それぞれの包含関係、隣接関係が抽出される。

【0078】この関係を木構造で示したものを本発明におけるレイアウト構造と呼ぶ。図6に示した例では、フレーム画像内には「段」の数は1である。また、各フレーム画像において含まれる文字列の数は3である。各文字列に含まれる文字の数は7～8である。

【0079】以上の包含関係と文字列どうし、または文字どうしの隣接関係を木構造を用いて表現したのが、図7(d-1)～(d-3)である。同木構造において、各ノードは画像内における物体のフレーム画像上での位置の情報を格納している。また、各ノードの格納順序は、横書きの場合、左から右、上から下の順に、文書を読み取る順序に合わせており、ノードの格納順序で画像中の物体の隣接関係が表現される。

【0080】レイアウト構造対応づけ部41での以降は第1の実施の形態における構造情報対応づけ部40において構造情報を対応づけるのと同様に物体の対応づけ、フレーム画像間の位置ずれ検出が行われる。図7(e)は同図(d-2)と(d-3)の間で行われるレイアウト構造の対応づけを示したものである。この結果として、複数枚のフレーム画像から生成されたレイアウト構造の木構造情報が統合され、レイアウト構造記憶部51に出力される。また、フレーム画像間の位置ずれ量も同時に出力される。以降の処理は本発明の第1の実施の形態と同様に行われる。

【0081】以上のように、本発明の第2の実施の形態

によれば、入力対象を文書としてフレーム画像の構造解析をレイアウト解析により行っている。レイアウト構造は印刷文書において安定した構造特徴であり、高速に構造情報の抽出が可能である。このため、フレーム画像の位置ずれを検出するための相関計算の回数を削減できるとともに、位置ずれ量の検出精度を安定に向上させることができる。

【0082】本発明の第3の実施の形態について図8を参照して詳細に説明する。

【0083】本発明の第3の実施の形態の構成は、第1の実施の形態の構成に加えて、位置ずれ残差検出部70と超解像処理部80とを含んで構成される点で異なる。

【0084】位置ずれ残差検出部70は、構造情報記憶部50に記憶されているフレーム画像間の位置ずれ量、構造情報と、それらの算出に用いられた複数のフレーム画像を入力し、対応する物体の画像に関して画素の大きさ未満の精度で位置ずれ量を算出する。これを以降でサブピクセル位置ずれ量と呼ぶ。サブピクセル位置ずれ量は超解像処理部80に出力される。

【0085】超解像処理部80は、フレーム画像間の位置ずれ量、構造情報、フレーム画像、サブピクセル位置ずれ量を入力する。フレーム画像間の位置ずれ量およびサブピクセル位置ずれ量を考慮してフレーム画像どうしを重畳し、画素の間を各フレーム画像の輝度値に基づき内挿することで、解像度を改善した画像を生成する。

【0086】本発明の第3の実施の形態の動作を図9を参照して詳細に説明する。

【0087】入力された各フレーム画像の構造情報に対応づけ、記憶するまでの処理は本発明の第1の実施の形態の動作と同じであるためここでは省略する。

【0088】複数のフレーム画像から得られた構造情報を統合して生成された構造情報に基づき、各フレーム画像から対応する個々の物体の画像を取り出す（図9（a-1）（a-2））。取り出された物体は所望の解像度になるように拡大され、位置ずれ残差検出部70において、再度位置ずれ量を検出し、サブピクセル位置ずれ量とする。サブピクセル位置ずれ量の検出は構造情報の対応づけを行った方法と同様に、拡大された物体の画像どうしを平行移動しながら重畳し、距離値が最も小さくなったときの移動量を位置ずれ量とする。

【0089】超解像処理部80では、位置ずれ残差検出部70で検出されたサブピクセル位置ずれ量を考慮して、各フレーム画像の対応する物体の拡大画像を重ね合わせ、画素値の内挿処理を行い、物体ごとの高解像度画像を生成する。これを超解像画像と呼ぶ。

【0090】内挿の方法としては、例えば信学技報、I E 90-54、23ページから28ページに記載された方法を用いることができる。

【0091】画像合成部60では、各物体に対して得られた超解像画像を構造情報に基づいて配置して全体の画

像を生成する。

【0092】次に、図9を用いて、本発明の第3の実施の形態により超解像画像を生成する過程を説明する。

【0093】第1の実施の形態の説明で用いた入力対象である図2を用いる。図2の各フレーム画像を構造解析し、図9（a-1）（a-2）に示すようにそれぞれのフレーム画像から対応する物体を取り出す。それぞれの取り出した物体を拡大すると、同図（b-1）（b-2）のようになる。拡大された画像どうしでサブピクセル位置ずれ量を検出し、これに基づき両画像を重畳し、内挿輝度値を内挿することで、図9（c）のような、解像度が改善された超解像画像を生成する。以上の処理を画像内の全物体に対して行い画像全域にわたる高解像度画像を生成する。

【0094】以上のように、本発明の第3の実施の形態によれば、内挿による解像度の改善処理の効果の不安定さを補い、安定に解像度の向上を実現できる。

【0095】本発明の第4の実施の形態について図10を参照して詳細に説明する。

【0096】第4の実施の形態では、第2の実施の形態と同様に入力する対象を文書に限定する。本発明の第4の実施の形態と第2の実施の形態の構成を図10と図5で比較すると、第4の実施の形態は第2の実施の形態に加えて位置ずれ残差検出部70と超解像処理部80とを含んで構成される点で異なる。

【0097】位置ずれ残差検出部70と超解像処理部80は第3の実施の形態の構成と同様であるため省略する。

【0098】本発明の第3の実施の形態の動作を図10を参照して詳細に説明する。

【0099】入力された各フレーム画像のレイアウト構造に対応づけ、記憶するまでの処理は本発明の第2の実施の形態の動作と同じであるためここでは省略する。

【0100】位置ずれ残差検出部70では対応づけられた文字を両フレーム画像から取りだし、第3の実施の形態と同様にサブピクセル位置ずれ量を検出する。

【0101】超解像処理部80では、拡大された各文字画像をサブピクセル位置ずれ量だけ平行移動して重畳した後、内挿処理を行い、超解像画像を生成する。

【0102】次に、図9を用いて、本発明の第4の実施の形態の一実施例の動作を詳細に説明する。

【0103】図2に示したものと同様の入力対象を用いる。図2の各フレーム画像を構造解析し、図9（a-1）（a-2）に示すようにそれぞれのフレーム画像から対応する物体を取り出す。（図9では「○」で示されているが、本実施の形態の入力対象は文書であるので、この「○」は本実施例では文字一般を表わす。）それぞれの取り出した物体を拡大すると、同図（b-1）（b-2）のようになる。拡大された画像どうしでサブピクセル位置ずれ量を検出し、これに基づき両画像を重畳

し、内挿輝度値を内挿することで、図9(c)のような、解像度が改善された超解像画像を生成する。

【0104】以上の処理を画像内の全物体に対して行い画像全域にわたる高解像度画像を生成する。

【0105】以上のように、本発明の第4の実施の形態によれば、本発明の第3の実施の形態に加えて、入力対象を文書とし、フレーム画像の構造解析をレイアウト解析により行っている。レイアウト構造は印刷文書において安定した構造特徴であり、高速に構造情報の抽出が可能である。このため、フレーム画像の位置ずれを検出するための相関計算の回数を削減できるとともに、位置ずれ量の検出精度を安定に向上させることができる。

【0106】また、予めレイアウト構造を用いてフレーム画像間の位置ずれ量が画素のサイズの精度で検出されているため、超解像処理のためのサブピクセル位置ずれ量の検出を画像全体に対して行う必要がなく、高速化が実現される。

【0107】更に、文書には多くの空白が含まれるが、それらを処理の対象から外すことができ、高速な超解像処理が可能となる。

【0108】本発明の第5の実施の形態について図11を参照して詳細に説明する。

【0109】第5の実施の形態では、第4の実施の形態と同様に文書を処理対象とし、文書内に書かれた文字列の認識を目的とする場合に限定する。図11を参照すると、本発明の第4の実施の形態の構成は、第2の実施の形態の構成に加えて、個別文字認識部90と、文字認識結果結合部61とを含んで構成される点で異なる。

【0110】本発明の第5の実施の形態の動作を詳細に説明する。

【0111】個々の文字画像について超解像処理部80までの処理により解像度が改善された文字の超解像画像を生成する部分までは第4の実施の形態の構成と同様であるため省略する。

【0112】個別文字認識部90は、超解像処理部80の出力する文字ごとの超解像画像を入力し、各文字を対応する文字コードに変換する。文字認識の方法としては、例えば、信学技報PRU90-20に記載された方法を用いることができる。

【0113】文字認識結果結合部61は、個別文字認識部90の出力する文字認識結果の文字コードと、レイアウト構造対応づけ部41の出力する統合されたレイアウト構造を入力し、文書全面を文字コードに変換したテキスト情報を出力する。統合されたレイアウト構造は文字画像がどのフレームのどの位置に書かれていたものかを表現しているため、これに従って文字コードを配置し直す。

【0114】次に、図6及び図7に示したものと同様の文書を入力対象として用いた場合の本発明の第5の実施の形態の動作を説明する。

【0115】本発明の第4の実施の形態と同様な処理により、解像度が改善された超解像文字画像を生成する。これらはそれぞれ個別文字認識部90に入力され、それぞれ個別文字認識処理によりそれぞれ対応する文字コードに変換される。変換された文字コードをレイアウト構造対応づけ部41において統合されたレイアウト構造にしたがって図12のように文書の読み取り順に文字コードを配置し直し、文書の認識結果とする。

【0116】以上のように、本発明の本発明の第5の実施の形態によれば、統合したレイアウト構造を用いて文字認識結果の文字コードを配置し、文書の認識結果とする。このため、本発明の第4の実施の形態の効果に加えて、文字認識処理を文書の全面的画像を再構成することなく、広範囲にわたる文書の認識を少ないメモリで行うことができる。

【0117】また、文字認識処理は、本手法で用いている他の処理に比較して、一般に大量の処理時間を必要とする。このため、個々のフレーム画像を高解像度取得し、各フレーム画像内の文字画像をそれぞれ文字認識し、文字列の接続を行う場合に比較して、処理時間を短縮することができる。

【0118】

【実施例】図13は本発明の第1の実施例である。ここではまず第1の実施の形態に基づく実施例を説明し、その後これを元に第2、3、4及び5に基づく実施例を説明する。

【0119】カメラ300は例えばデジタルスチルカメラ、デジタルビデオカメラ、CCDカメラなどである。カメラ300を用いて撮像対象を走査するが、走査は一定の方向、速さで行う必要がなく、手動で行うことができる。

【0120】走査して得られる動画画像または分割静止画像は複数枚のフレーム画像からなり、これらフレーム画像はビデオキャプチャボード301を介してパーソナルコンピュータ302に取り込まれる。

【0121】取り込まれたフレーム画像は画像記憶部20としての外部記憶装置303に記憶される一方、構造解析部30として働くパーソナルコンピュータ302により図3のような構造情報を抽出され、各フレーム画像の構造情報は構造情報記憶部50として働く外部記憶装置303に記憶される。

【0122】構造情報対応づけ部40として働くパーソナルコンピュータ302は、同一物体を含むフレーム画像の構造情報を統合して図4のような構造情報を生成する。

【0123】画像合成部60として働くパーソナルコンピュータ302は、統合して得られた構造情報と、画像記憶部20として働く外部記憶装置303に記憶されたフレーム画像をもとに、フレーム画像を合成して合成画像を得る。



【0124】こうして得られた合成画像は、ディスプレイ装置、プリンタなどの出力装置 304 により表示される。

【0125】第 2、4、5 の実施の形態に基づく実施例においては、撮像対象が文書に限定されるため、上記の説明で構造解析部 30、構造情報対応づけ部 40、構造情報記憶部 50 はそれぞれレイアウト構造解析部 31、レイアウト構造対応づけ部 41、レイアウト構造記憶部 51 に置き換えられる。

【0126】第 3、4、5 の実施の形態に基づく実施例においては、超解像画像の挿入が行われるため、上記に付け加えて、パーソナルコンピュータ 302 は位置ずれ残差検出部 70 および超解像処理部 80 として機能する。

【0127】第 5 の実施の形態に基づく実施例においては、超解像処理部 80 として機能したパーソナルコンピュータ 302 により、文書中の各文字についての超解像画像を生成し、これを個別文字認識部 90 として働くパーソナルコンピュータ 302 により対応する文字コードに変換する。文字認識結果結合部 61 として働くパーソナルコンピュータ 302 により、文字コードとレイアウト情報から撮像対象の文書が復元され、出力装置 304 から出力される。

【0128】図 14 は本発明の第 2 の実施例である。第 1 と第 2 の実施例の違いは、カメラ 300 とビデオキャプチャボード 301 の間にビデオデッキ、DVD (デジタルビデオディスク) 等の動画像記録装置 305 が接続されていることである。このため本実施例はリアルタイムの画像だけではなく、録画された画像についても処理を行うことが出来る。

【0129】

【発明の効果】第 1 の効果は、フレーム画像の位置ずれを検出するための相関計算の回数を削減でき、位置ずれ量の検出精度の向上も実現できる。

【0130】その理由は、入力された各フレーム画像の構造を解析し構造情報どうしの対応づけを行うためである。

【0131】第 2 の効果は、フレーム画像の位置ずれを検出するための相関計算の回数を削減できるとともに、位置ずれ量の検出精度が劣化しやすい文書画像に対しても安定に検出精度を向上させることができる。

【0132】その理由は、レイアウト構造は印刷文書において安定した構造特徴であり、高速に構造情報の抽出が可能であるためである。

【0133】第 3 の効果は、超解像処理のためのサブピクセル位置ずれ量の検出を画像全体に対して行う必要がなく、高速化が実現される。高速な超解像処理が可能となる。

【0134】その理由は、予めレイアウト構造を用いてフレーム画像間の位置ずれ量が画素のサイズの精度で検

出されているためである。また、文書には多くの空白が含まれるが、効率よくそれらを処理の対象から外すことができるためである。

【0135】第 4 の効果は、文書画像に対する解像度改善処理の高速化、高精度化が実現される。

【0136】その理由は、予めレイアウト構造を用いてフレーム画像間の位置ずれ量が画素のサイズの精度で検出されているため、超解像処理のためのサブピクセル位置ずれ量の検出を画像全体に対して行う必要がなく、また、文書には多くの空白が含まれるが、それらを処理の対象から外すことができるためである。

【0137】第 5 の効果は、広範囲にわたる文書の認識を少ないメモリで行うことができることである。また、各フレーム画像内の文字画像をそれぞれ文字認識し、文字列の接続を行う場合に比較して、処理時間を短縮することができる。

【0138】その理由は、統合したレイアウト構造を用いて文字認識結果の文字コードを配置し、文書の認識結果とし、文字認識処理を文書の全面の画像を再構成することなく文書の認識処理が行えるためである。

【図面の簡単な説明】

【図 1】本発明の第 1 の実施の形態のブロック図である。

【図 2】本発明の第 1 の実施の形態によるの画像の合成を表す説明図である。

【図 3】構造情報の例である。

【図 4】図 3 の 2 つの構造情報を変換して得た構造情報である。

【図 5】本発明の第 2 の実施の形態のブロック図である。

【図 6】フレーム画像からレイアウト情報を抽出する過程を表わす図である。

【図 7】図 6 で抽出されたレイアウト情報を統合する過程を表わす図である。

【図 8】本発明の第 3 の実施の形態のブロック図である。

【図 9】超解像画像を生成する過程を説明する図である。

【図 10】本発明の第 4 の実施の形態のブロック図である。

【図 11】本発明の第 5 の実施の形態のブロック図である。

【図 12】本発明の第 5 の実施の形態の動作を表わす説明図である。

【図 13】本発明の第 1 の実施例のブロック図である。

【図 14】本発明の第 2 の実施例のブロック図である。

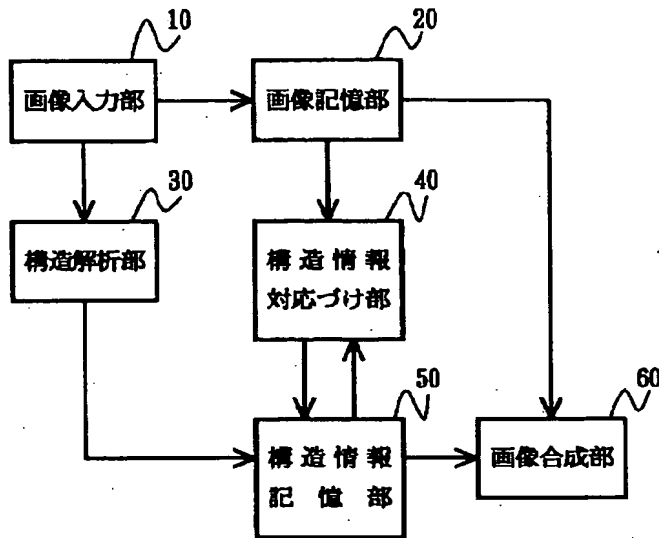
【符号の説明】

10 画像入力部  
20 画像記憶部  
30 構造解析部

- 31 レイアウト構造解析部
- 40 構造情報対応づけ部
- 41 レイアウト構造対応づけ部
- 50 構造情報記憶部
- 51 レイアウト構造記憶部

- 60 画像合成部
- 61 文字認識結果結合部
- 70 位置ずれ残差検出部
- 80 超解像処理部
- 90 個別文字認識部

【図1】



【図3】

## (1フレーム目)

物体1 101: 形状「△」、面積S101、回転角0度  
 物体2との相対位置 $\gamma_{101}$   
 物体3との相対位置 $\gamma_{103}$   
 位置 $x_{101}$

物体2 102: 形状「□」、面積S102、回転角0度  
 物体1との相対位置 $\gamma_{101}$   
 物体3との相対位置 $\gamma_{102}$   
 位置 $x_{102}$

物体3 103: 形状「○」、面積S103、回転角0度  
 物体1との相対位置 $\gamma_{102}$   
 物体2との相対位置 $\gamma_{103}$   
 位置 $x_{103}$

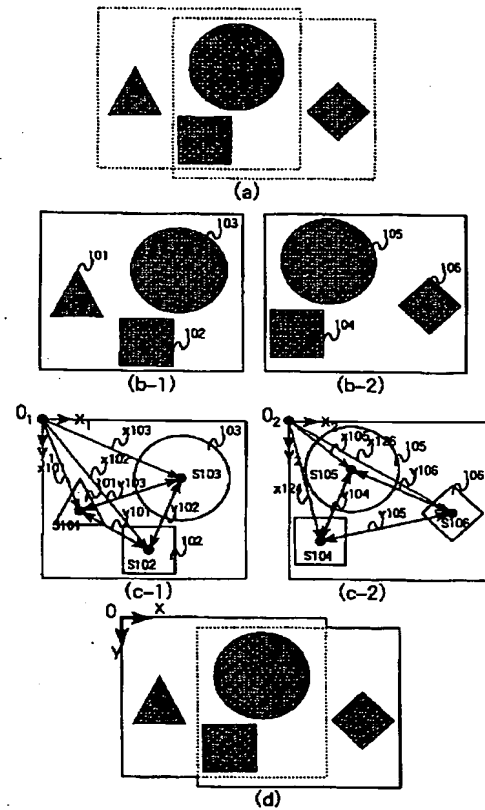
## (2フレーム目)

物体4 104: 形状「□」、面積S104、回転角0度  
 物体5との相対位置 $\gamma_{104}$   
 物体6との相対位置 $\gamma_{106}$   
 位置 $x_{104}$

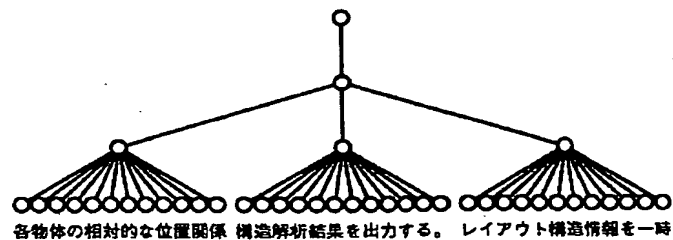
物体5 105: 形状「○」、面積S105、回転角0度  
 物体4との相対位置 $\gamma_{104}$   
 物体6との相対位置 $\gamma_{105}$   
 位置 $x_{105}$

物体6 106: 形状「□」、面積S106、回転角45度  
 物体4との相対位置 $\gamma_{105}$   
 物体5との相対位置 $\gamma_{106}$   
 位置 $x_{106}$

【図2】



【図12】



【図4】

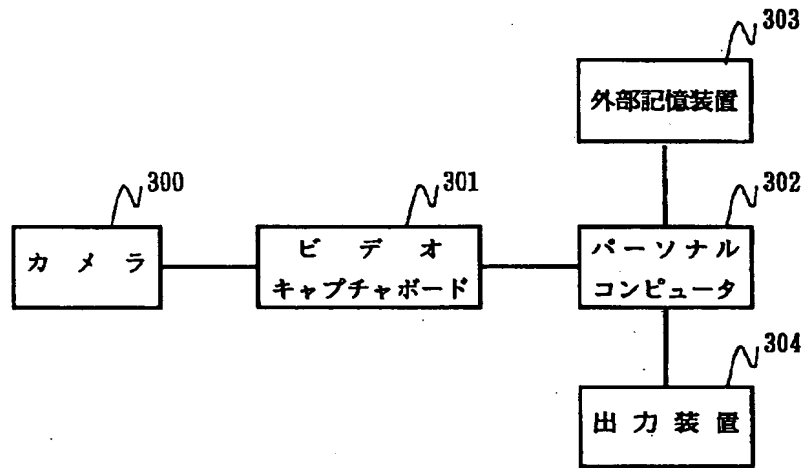
物体1 101:形状「△」、面積S101、回転角0度  
 物体2&4との相対位置 $\psi$ 101  
 物体3&5との相対位置 $\psi$ 103  
 位置 $x$ 101

物体2&4 102:形状「□」、面積S102、回転角0度  
 物体1との相対位置 $\psi$ 101  
 物体3&5との相対位置 $\psi$ 102  
 物体6との相対位置 $\psi$ 105  
 位置 $x$ 102

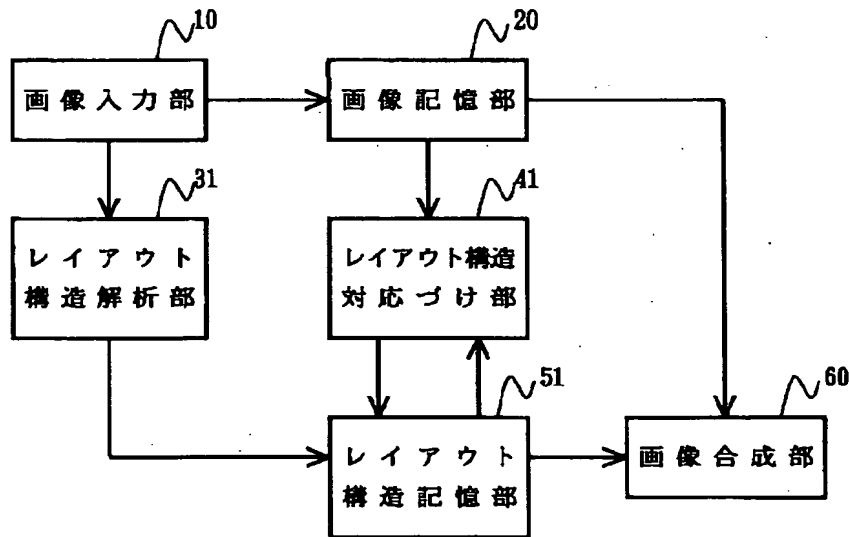
物体3&5 103:形状「○」、面積S103、回転角0度  
 物体2&4との相対位置 $\psi$ 102  
 物体1との相対位置 $\psi$ 103  
 物体6との相対位置 $\psi$ 105  
 位置 $x$ 103

物体6 106:形状「□」、面積S106、回転角45度  
 物体2&4との相対位置 $\psi$ 103  
 物体3&5との相対位置 $\psi$ 106  
 位置 $x$ 106-a

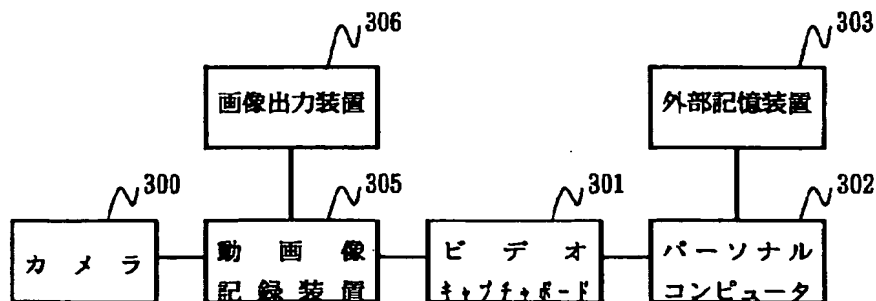
【図13】



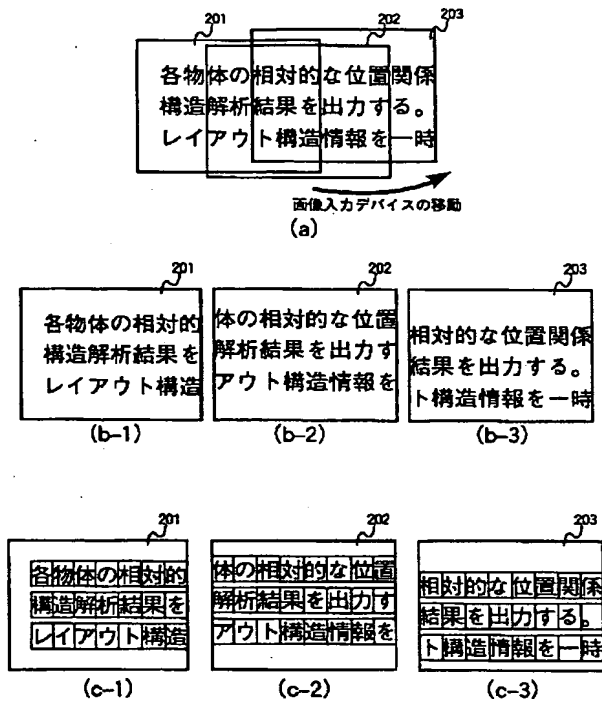
【図5】



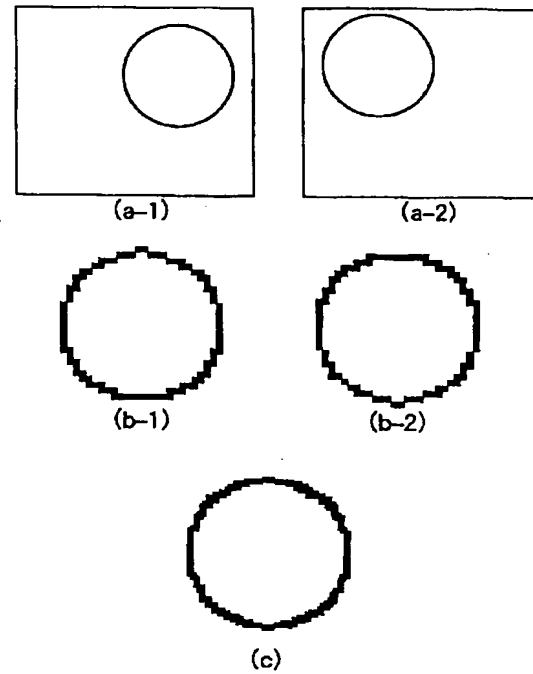
【図14】



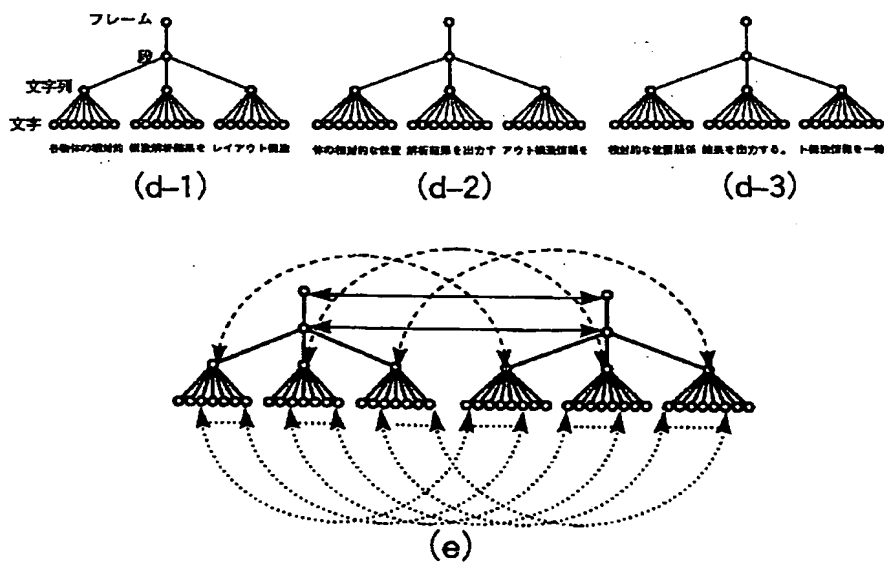
【図6】



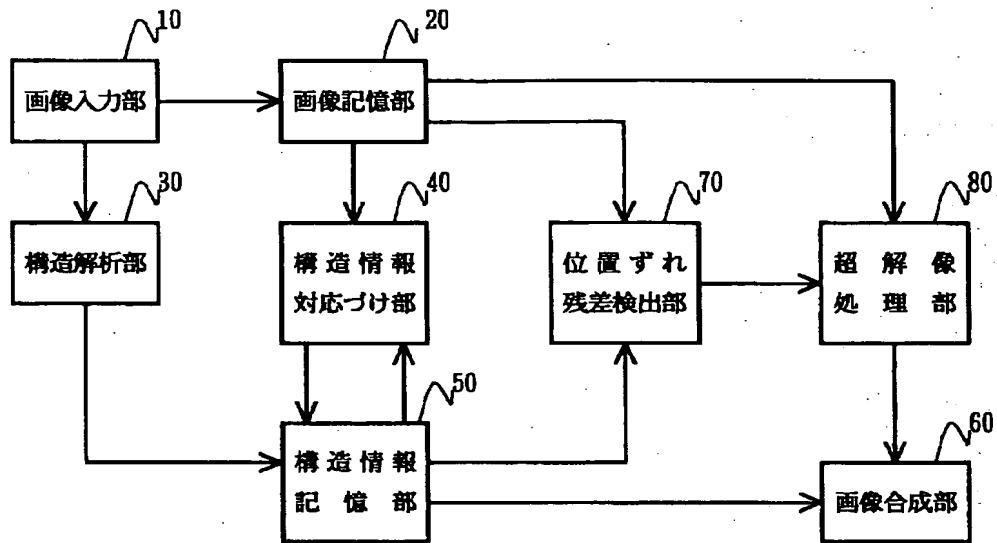
【図9】



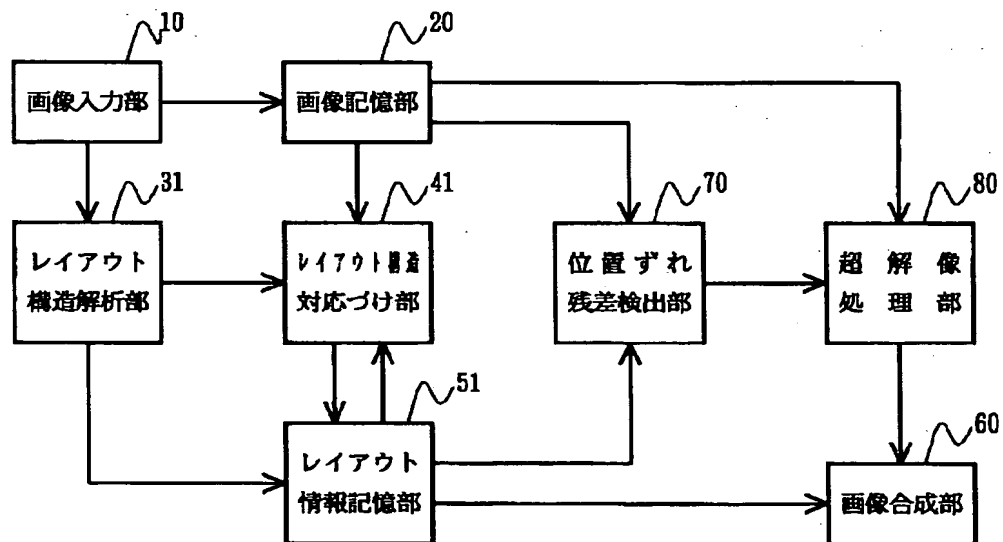
【図7】



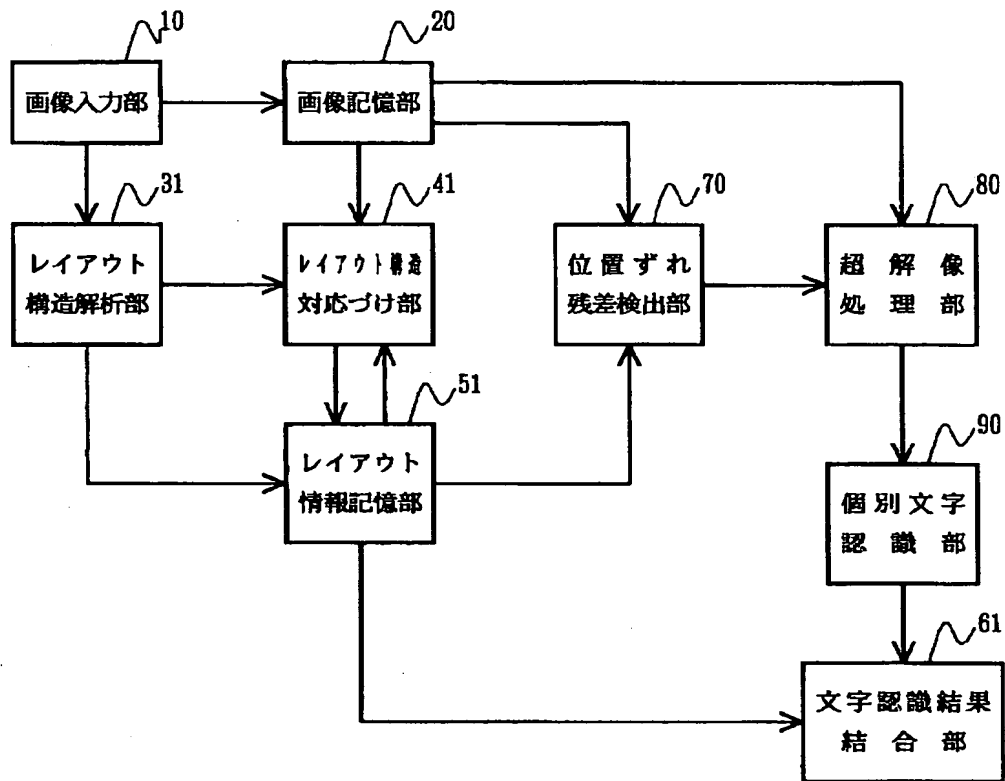
【図8】



【図10】



【図11】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**